

1 MSSKTASTNS IAQARRTVQQ LRLEASIERI KVSASADLM SYCEEHARSD
51 PLLMGIPTSE NPFKDKKTCI IL*

51 PLLMG IPTSE NPFKDKKTCI IL*

Figure 2: cDNA sequence of mGy12 variant 1 (SEQ ID NO:2)

1 CTAGAATTCA GCGGCCGCTG AATTCTAGGC GACGACGGCG AAGAGTGAGT
51 GCCAAGGTTC ATATGGGAAG GACTTTGGGG TGAGCATCTT CTCTATTTCC
101 AGCTGGCTTT TCTGATTTTC AGAAAGAAGA CTCATCAAAG ATGTCCAGCA
151 AGACGGCAAG CACCAACAGC ATAGCCCAAG CCAGGAGAAC TGTGCAGCAG
201 CTGAGATTGG AAGCCTCCAT CGAAAGAATA AAGGTCTCAA AAGCATCAGC
251 AGACCTGATG TCATACTGTG AGGAGCATGC CCGGAGCGAC CCCCTGCTGA
301 TGGGCATACC GACCTCAGAA AACCCGTTCA AGGATAAGAA GACCTGCATC
351 ATCTTATAGT GGACCAGGAA GCGCCCCTTG CCTCTTAACG CAAACCACAG
401 CAGCAACCTG AAGGGATTCC TTCAGCTTAC CTGGTAACCA CAGCTAGTAA
451 CTAAACACC CTTCTCTCGG AATAATAGAC CCTGAAGTCT CTCTTTTTCA
501 AGTTGTCCTT TCTTCACCCT TTACTGATTT AATACAGAAT AACAATCTTA
551 TTTTCTATTT GATAACTATG GTATCATATT GGGTTACTGT ATAAGGAAAA
601 TGGCAGGGGA GTTGTGGGAA GCTTGTCTTT ACAAATATA ATTGATTAAG
651 ATATGTCAAG ACCTACATTG TCTAAGCACC GGCAAATTAA AATGTCGAGA
701 ATCACTTCAG TCAAAAACCT TTATATTCTG TTCTTAATAA TGTTTGTGCC
751 AACCTATATC CCATGTAAGG GATCTGGGGA GGAGGCATGT GTCTACAACC
801 ATACCTTTTT GCACTATGGG CACTAACCAC CCTGAACTT CCTGCGGTAG
851 CTCCCTCCCT TCAGAGTTAC ATCATTATCC TGA CTCTGTG TAGGTAAATT
901 TCCGTGAAAT TTTTGTACAA AAAAAAGGTA ATGAAAGAAC GTTGCAAAGA
951 TCATCTGCAT TATAATGAGT TGATGCTGTT CTCCTCCTC TCTTGGAATT
1001 GTGCTGGCCC CTTAGTCTAC AATAAACTGT GCCAATTAAA AACCTAAGGC
1051 TAAAACTGAA AGCCCTTTGA TGGGGTCTTA ACTCATATCA GTCATTTGGG
1101 CTTCTCTGAT CCTGAGGCTA AGAAAGGGGA AGAGACCCTC AGGAGGCAGC
1151 TTCCACTCCA GGGCTCTTGA TCTCTGCTGG ATTGGGGGTG GCCACCTCAG
1201 AACTTCCAC CCTCATGACT GGAATGGAAG AGGGGACCGA GAGCCTCACA

1251 ATCTCGGAGA GGGAGGAGAA ATTCTTAAAA ACAGCTGCTC TCCTGCGCCC
1301 AGCTTCACAG GCAGCCCTGC CCCTTTCTCC TCACCAGCAT GGTACCTGCC
1351 CTTACTGCTA GAGCAGCTGC TTGTAGAGGG ACATTCCCTC CTTCCCAGTT
1401 TTAAGTGGTG GACCACAGTG GGGGGAAAAA CATTCAAGCG ATATAAAGAC
1451 ACTTGGGCTC TTTGCAGATG CCTATACTTC CAACACTACC ATGTCCACAA
1501 ACCACCCTGG GGGAGGGCCC TTCCAAAGGG AGGCTTGCTA GTTTCAGCGT
1551 CTAGCAGTTG GGTCTCACT TTTACTCCAA TTGTGAAAAT AGCCCACGTA
1601 CCCTCGCAGT GTCCAGTAGG GATCCCAGAG GCACATAACC AAGAAAGGAT
1651 TTTGACTTTG TCACAGTGAC TATTTAAAAT AATCTATTCG AAGTCCAAAC
1701 CAAACACAAA GCCTGTGATA TTTTAGGTTA TTAAGGTAAC TGCTAATGAA
1751 GGATTTTAAA AAGTGTCTCT AAAAAGGACT TAGCCCCGGG AGTTGTTTAT
1801 AAAATTTCCC CCACTTGTAT ACAGTGTGCA CTAAAAGAAA ATGTATTTTA
1851 ATATCTAATG CCTGGGCTCT GAGCGTCATG CTTCTTGGTG GTAAACATGC
1901 AGTCCTGTTC CTAAGTGACT CAGAGGCATC AGAATTTCTC CACGTTACCC
1951 ATCTGCTTGG CACTCGGAAC TGAGCGTGTG AAATCCATAG CGCTGCCCAC
2001 AACCTGTTCT CACTGCTTAG CTCCCAGCTG GATTAAAGAC ACCTGCTCAG
2051 GCGGGAGAGA GAGAGAGAGA GCGAGCTTTT ACCTTGAAA AGGTAAAGAT
2101 GGAAATGTAC ACCAAAAAAG ACAATTTTTA CATTAAATGG AACATTCTTT
2151 TTTTTTACAA GTATATTTTT CTACTGATAG TTTCAGAACA CTAATCTTAT
2201 ATTCACTCTA ATCTTAAACA TGTTTCTTTA AATATTTATA AGGCAGTTTA
2251 TTACAGAATA TTTTCATGCA ATCATGTGCA CATTATTGGT AGCAAACATA
2301 GTATATCCTT TAGTACTTTA GCATATTTTT GTTAAAATAC TTTTAATGGT
2351 AAGAAATGAA CTTGAGGTCC CAGGAGGTTT TGTTGCCTTT TCATTGATTA
2401 GAGACAATAA ATATCTTGTA ACTTCCTAAC CAGATCTGAG CTGTGCTCAC
2451 AATAATAATA ATGAAATCAG ATTCTTTGAT GCTGGACTCA GGAGGGAAAT

2501 CATTAGCCAA CTGTTGACTT ACTTATAGCT AGATGTCTAT GTGAGAAAGT
2551 ATAATATATA TATATACACA TATATATGAC ATGTAAGAGT CACTTTTATT
2601 TATCTGTCTT TGTTCACTTA TGAAGCCGGT AACTGCAGCA GTATGTTGGT
2651 GATGTCATGA TGCACAGAAG TCCCATGTGG AGTGTTTTTC CCACACTGAC
2701 AACTTGGCCT CCTTTCTGTG TGTTCACTCT GTTGTCTGAA CTAACACTCC
2751 CGCGAGCACT ATACTCTTTA TACTCTGATC CCCCTAGTTC ATCTTAAATT
2801 TGTCTGTGGC CCTGGCAAGA TAGCGTACAC AAGATTCCAT GACTCCAGAG
2851 CATCTTGAAG AAACATACAT ATTTTGAAAG AGGGGAAATG TAGCAGATAG
2901 TTCACAAGCT GCGGGTTGTA GCTAAATATT CCATTTCTTT GAAATCATGT
2951 TTCTAAATTC TTTACCATCA GAAAGAAAAG GAGTGTCTATA CACTTTCAAG
3001 GGAAGGCTTG GTCTGCGTTT TCTGTGTTTG GATTATTTTT ATACTTTGCT
3051 GATCTTTAAG CTATCCATGG GGGAAATTTT ATACCAACGA GTTAATAATT
3101 CTCATTCATC GTTTACACAA TGTAACATGT GTCATACTGG GGCCAGCGAG
3151 ATGGCTCAGT AGGTAAAGGT GCTTGATGCT AAGCCCGGCA GCCTGTGTTT
3201 CATCTACAGG ATGCACAACA TAAAAGAAAA GATCTGATTC CCACAGGTTC
3251 TCTTCTGACC TACACACACA CACACTAAAA TAACATTTAA AAATATGTGC
3301 CAAATTATAT TTGTTTCGGGT GCCACCTTCC ACCAGCTTAC CACTACGGTA
3351 GAACTGTCAA ATTCATCTCC CTGAATTTGT CTTAAAGGGG TGTCCATGCA
3401 CAGGCCCAAG AGTCACCTCC AATGAAATAA ATGTAATACT GAAGTATGCC
3451 ATGATGTTTG TTGTTTTCTT TCATCGTAAG CCTGTAAGCA GGAAAAATAC
3501 GTCAAATCAG ATAGAATAGA GCATTTACCA GTGGTCGATG GCCTGGTCAG
3551 TCCTGTGCCG GGTGACTTAG GACCAGGCAC GTCAGCTCTC TGAGCCTCCC
3601 CTTCCCTTGT TGTCACAAGG GAATAGAAGC AGAAGAAGCT GAGAGCCTCC
3651 CTATTCCCAG ATGCCCTGGT GGAATGACCT GCCTCTCTGC CGTTTCTGCC
3701 AACGTGTTGG TGCTATAAGC TGCTTCAAAT ACCAGTTTGT CTGTAGTGTG
3751 TACTCACCTA ATCACTTGTT ATCCAGTGCC TGTCTAGGT TTATGGACTT

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(36)	(37)	(38)	(39)	(40)	(41)	(42)	(43)	(44)	(45)	(46)	(47)	(48)	(49)	(50)	(51)	(52)	(53)	(54)	(55)	(56)	(57)	(58)	(59)	(60)	(61)	(62)	(63)	(64)	(65)	(66)	(67)	(68)	(69)	(70)	(71)	(72)	(73)	(74)	(75)	(76)	(77)	(78)	(79)	(80)	(81)	(82)	(83)	(84)	(85)	(86)	(87)	(88)	(89)	(90)	(91)	(92)	(93)	(94)	(95)	(96)	(97)	(98)	(99)	(100)
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	

Figure 3: cDNA sequence of mGy12 variant 2 (SEQ ID NO:3)

1 GCAGCGGCGG CGGCGGCGAC GACGGCGAAG AGTTCATATG GGAAGGACTT
51 TGGGGTGAGC ATCTTCTCTA TTTCCAGCTG GCTTTTCTGA TTCACCCAC
101 CATTTAAAAC CTGGAGGCAC TGGGCCACAC AAAGCCTTGC TGATTTTCAG
151 AAAGAAGACT CATCAAAGAT GTCCAGCAAG ACGGCAAGCA CCAACAGCAT
201 AGCCCAAGCC AGGAGAACTG TGCAGCAGCT GAGATTGGAA GCCTCCATCG
251 AAAGAATAAA GGTCTCAAAA GCATCAGCAG ACCTGATGTC ATACTGTGAG
301 GAGCATGCCC GGAGCGACCC CCTGCTGATG GGCATACCGA CCTCAGAAAA
351 CCCGTTCAAG GATAAGAAGA CCTGCATCAT CTTATAGTGG ACCAGGAAGC
401 GCCCCTTGCC TCTTAACGCA AACCACAGCA GCAACCTGAA GGGATTCCTT
451 CAGCTTACCT GGTAACCACA GCTAGTAACT AAAACACCCT TCTCTCGGAA
501 TAATAGACCC TGAAGTCTCT CTTTTTCAAG TTGTCCTTTC TTCACCCTTT
551 ACTGATTTAA TACAGAATAA CAATCTTATT TTCTATTTGA TAACTATGGT
601 ATCATATTGG GTTACTGTAT AAGGAAAATG GCAGGGGAGT TGTGGGAAGC
651 TTGTCTTTAC AAAATATAAT TGATTAAGAT ATGTCAAGAC CTACATTGTC
701 TAAGCACCGG CAAATTAAAA TGTCGAGAAT CACTTCAGTC AAAAACCTTT
751 ATATTCTGTT CTTAATAATG TTTGTGCCAA CCTATATCCC ATGTAAGGGA
801 TCTGGGGAGG AGGCATGTGT CTACAACCAT ACCTTTTTGC ACTATGGGCA
851 CTAACCACCC TGAAACTTCC TCGGGTAGCT CCCTCCCTTC AGAGTTACAT
901 CATTATCCTG ACTCTGTGTA GGTAAATTTT CGTGAAATTT TTGTACAAAA
951 AAAAGGTAAT GAAAGAACGT TGCAAAGATC ATCTGCATTA TAATGAGTTG
1001 ATGCTGTTCT CACTCCTCTC TTGGAATTGT GCTGGCCCCT TAGTCTACAA
1051 TAACTGTGTC CAATTAAAAA CCTAAGGCTA AACTGAAAG CCCTTTGATG
1101 GGGTCTTAAC TCATATCAGT CATTTGGGCT TCTCTGATCC TGAGGCTAAG
1151 AAAGGGGAAG AGACCCTCAG GAGGCAGCTT CCACTCCAGG GCTCTTGATC

1201 TCTGCTGGAT TGGGGGTGGC CACCTCAGAA ACTTCCACCC TCATGACTGG
1251 AATGGAAGAG GGGACCGAGA GCCTCACAAT CTCGGAGAGG GAGGAGAAAT
1301 TCTTAAAAAC AGCTGCTCTC CTGCGCCCAG CTTACAGGC AGCCCTGCCC
1351 CTTTCTCCTC ACCAGCATGG TACCTGCCCT TACTGCTAGA GCAGCTGCTT
1401 GTAGAGGGAC ATTCCCTCCT TCCCAGTTTT AACTGGTGGA CCACAGTGGG
1451 GGGAAAAACA TTCAAGCGAT ATAAAGACAC TTGGGCTCTT TGCAGATGCC
1501 TATACTTCCA ACACTACCAT GTCCACAAAC CACCCTGGGG GAGGGCCCTT
1551 CCAAAGGGAG GCTTGCTAGT TTCAGCGTCT AGCAGTTGGG TCCTCACTTT
1601 TACTCCAATT GTGAAAATAG CCCACGTACC CTCGCAGTGT CCAGTAGGGA
1651 TCCCAGAGGC ACATAACCAA GAAAGGATTT TGACTIONTGT ACAGTGACTION
1701 TTTAAAATAA TCTATTGAA GTCCAAACCA AACACAAAGC CTGTGATATT
1751 TTAGGTTATT AAGGTAAGT CTAATGAAGG ATTTTAAAA GTGTCCTCAA
1801 AAAGGACTTA GCCCCGGGAG TTGTTTATAA AATTTCCCC ACTTGTATAC
1851 AGTGTGCACT AAAAGAAAAT GTATTTTAAT ATCTAATGCC TGGGCTCTGA
1901 GCGTCATGCT TCTTGGTGGT AAACATGCAG TCCTGTTCTT AAGTGACTION
1951 GAGGCATCAG AATTTCTCCA CGTTACCCAT CTGCTTGGCA CTCGGACTION
2001 AGCGTGTGAA ATCCATAGCG CTGCCCACAA CCTGTTCTCA CTGCTTAGCT
2051 CCCAGCTGGA TTAAAGACAC CTGCTCAGGC GGGAGAGAGA GAGAGAGAGC
2101 GAGCTTTTAC CTTGGAAAAG GTAAAGATGG AAATGTACAC CAAAAAGAC
2151 AATTTTACA TTTAATGGAA CATTCTTTTT TTTTACAAGT ATATTTTTCT
2201 ACTGATAGTT TCAGAACACT AATCTTATAT TCACTCTAAT CTAAACATG
2251 TTTCTTTAAA TATTTATAAG GCAGTTTATT ACAGAATATT TTCATGCAAT
2301 CATGTGCACA TTATTGGTAG CAAACATAGT ATATCCTTTA GTACTTTAGC
2351 ATATTTTTGT TAAATACTT TTAATGGTAA GAAATGAACT TGAGGTCCCA
2401 GGAGGTTTTG TTGCCTTTTC ATTGATTAGA GACAATAAAT ATCTTGTAAC

2451 TTCCTAACCA GATCTGAGCT GTGCTCACAA TAATAATAAT GAAATCAGAT
 2501 TCTTTGATGC TGGACTCAGG AGGGAAATCA TTAGCCAACT GTTGACTIONAC
 2551 TTATAGCTAG ATGTCTATGT GAGAAAGTAT AATATATATA TATACACATA
 2601 TATATGACAT GTAAGAGTCA CTTTTATTTA TCTGTCTTTG TTCACTTATG
 2651 AAGCCGGTAA CTGCAGCAGT ATGTTGGTGA TGTCATGATG CACAGAAGTC
 2701 CCATGTGGAG TGTTTTTCCC AACTGACAA CTTGGCCTCC TTTCTGTGTG
 2751 TTCAGTCTGT TGTCTGAACT AACACTCCCC CGAGCACTAT ACTCTTTATA
 2801 CTCTGATCCC CCTAGTTCAT CTAAATTTG TCTGTGGCCC TGGCAAGATA
 2851 GCGTACACAA GATTCCATGA CTCCAGAGCA TCTTGAAGAA ACATACATAT
 2901 TTTGAAAGAG GGGAAATGTA GCAGATAGTT CACAAGCTGC GGGTTGTAGC
 2951 TAAATATTCC ATTTCTTTGA AATCATGTTT CTAAATTCTT TACCATCAGA
 3001 AAGAAAAGGA GTGTCATACA CTTTCAAGGG AAGGCTTGGT CTGCGTTTTTC
 3051 TGTGTTTGGA TTATTTTTAT ACTTTGCTGA TCTTTAAGCT ATCCATGGGG
 3101 GAAATTTTAT ACCAACGAGT TAATAATTCT CATTCACTGT TTACACAATG
 3151 TAACATGTGT CATACTGGGG CCAGCGAGAT GGCTCAGTAG GTAAAGGTGC
 3201 TTGATGCTAA GCCCGGCAGC CTGTGTTTCA TCTACAGGAT GCACAACATA
 3251 AAAGAAAAGA TCTGATTCCC ACAGGTTCTC TTCTGACCTA CACACACACA
 3301 CACTAAAATA ACATTTAAAA ATATGTGCCA AATTATATTT GTTCGGGTGC
 3351 CACCTTCCAC CAGCTTACCA CTACGGTAGA ACTGTCAAAT TCATCTCCCT
 3401 GAATTTGTCT TAAAGGGGTG TCCATGCACA GGCCCAAGAG TCACCTCCAA
 3451 TGAAATAAAT GTAATACTGA AGTATGCCAT GATGTTTGTT GTTTTCTTTC
 3501 ATCGTAAGCC TGTAAGCAGG AAAAATACGT CAAATCAGAT AGAATAGAGC
 3551 ATTTACCAGT GGTCGATGGC CTGGTCAGTC CTGTGCCGGG TGACTIONAGGA
 3601 CCAGGCACGT CAGCTCTCTG AGCCTCCCCT TCCCTTGTTG TCACAAGGGA
 3651 ATAGAAGCAG AAGAAGCTGA GAGCCTCCCT ATTCCCAGAT GCCCTGGTGG
 3701 AATGACCTGC CTCTCTGCCG TTTCTGCCAA CGTGTTGGTG CTATAAGCTG

3751 CTTCAAATAC CAGTTTGTCT GTAGTGTGTA CTCACCTAAT CACTTGTTAT
3801 CCAGTGCCTG TTCTAGGTTT ATGGACTTAA CTATTTCTGT GATGTTTCAT
3851 TTTTAGCCAT GTTAACTCCT AACACATATT CTCTTATGTC TCAGTAAAGT
3901 TTCATTTGAT AAGTTGTTGA GATTCTGTGA TTTGATAATA TTCTTCGGCT
3951 GTCCATCCAG CATCTTAATC ACTTTAAAAC TGTGATTGTT ATTTGCAACT
4001 CTGTTCTTTG GAAAGAATAA AAGCATTTTT TTTCACTTGC TAACATGCTC
4051 ACAAATGTGA GAGAAGAGTC ATTAAAAGCT TTACTTACTG GGTCAGTGCG
4101 TCATTGACTC CTTTCTGTGT TTTGCCCAAT AAATTAATAA AAGACCAAAA
4151 AAAAAAAAAA AAAAAAAAAA AAAAA

1. MSSKTASTNN IAQARRTVQQ LRLEASIERI KVS KASADLM SYCEEHARSD
51. PLLIGIPTSE NPFKDKKTCI IL

1. MSSKTASTNN IAQARRTVQQ LRLEASIERI KVSASADLM SYCEEHARS
51. PLLIGIPTSE NPFKDKKTCI IL

1000 900 800 700 600 500 400 300 200 100 0



FIGURE 5